

ENGLISH REVIEW: Journal of English Education  
Volume 7, Issue 1, December 2018

p-ISSN 2301-7554, e-ISSN 2541-3643  
<https://journal.uniku.ac.id/index.php/ERJEE>

# CRITERIA USED BY PROFICIENT VS. LESS PROFICIENT EFL TEACHERS IN PRAGMATIC ASSESSMENT: THE CASE OF REQUEST SPEECH ACT

**Misagh Hosseini**

*English Department, Isfahan (Khorasgan) Branch, Islamic Azad University, Isfahan, Iran*  
E-mail: [mishoseini2000@gmail.com](mailto:mishoseini2000@gmail.com)

**Ehsan Rezvani**

*English Department, Isfahan (Khorasgan) Branch, Islamic Azad University, Isfahan, Iran*  
E-mail: [rezvani\\_ehsan\\_1982@yahoo.com](mailto:rezvani_ehsan_1982@yahoo.com)

APA Citation: Hosseini, M., & Rezvani, E. (2018). Criteria used by proficient vs. less proficient EFL teachers in pragmatic assessment: The case of request speech act. *English Review: Journal of English Education*, 7(1), 9-16. doi: 10.25134/erjee.v7i1.1529.

Received: 29-08-2018

Accepted: 30-10-2018

Published: 01-12-2018

**Abstract:** The issue of EFL teachers' rating criteria and patterns in Interlanguage pragmatics (ILP) assessment is new and needs rigorous analysis. The purpose of this study was to reveal important variables such as raters' criteria and rating patterns by analyzing the ILP assessment process of Iranian non-native English speaking raters (NNESRs) of both high and low proficiency levels based on the request speech act. The data for this study was collected through a discourse completion test (DCT) and a rating questionnaire from 40 Iranian EFL teachers and were later analyzed through descriptive analysis-test and chi-squares. The results showed that raters considered 9 criteria, including pragma-linguistic and socio-pragmatic components of language which raters noted differently through eight request situations. The results showed that raters considered nine criteria, including pragma linguistic and socio-pragmatic components of language which were noted differently through eight request situations. Among the considered criteria, the highest frequencies belonged to the criteria of authenticity, query of preparatory and softness, and interlocutors' relationship used by high proficiency teachers, whereas low proficiency teachers used the highest frequencies of accuracy, style, and directness. The result of the study can have important connotations for teachers to consider teaching L2 pragmatics in language classes and in teacher training courses.

**Keywords:** *Interlanguage pragmatics; request; speech act; pragmatic assessment.*

## INTRODUCTION

Interlanguage pragmatics (ILP) is regarded as the study of the development and use of strategies for linguistic action by non-native speakers. Incorporating L2 pragmatics into language teaching programs, especially in assessment area, plays a major role in second language process. Pragmatic assessment has been a salient theme in second language process and pragmatic instruction for years. According to Oller (1979, p. 39), "pragmatic proficiency test is any procedure or task that causes the learners to process sequences of elements in a language that conforms to the normal

contextual constraints of that language, and which requires the learners to relate sequences of linguistic element via pragmatic mapping to extra linguistic context."

The popularity of speech act studies in pragmatic is evident. Different categorizations associated with pragmatic studies are in both acquisitional areas, which deal with EFL learners' developmental issues, and comparative areas, which are dominantly of cross-cultural studies (Alcon-soler & Martinez- flor, 2008). Besides, pragma-linguistic and socio-pragmatic division of language knowledge (Leech &

Thomas, 1983) are another classifications in which linguistic and social dimensions of pragmatic knowledge are dealt with respectively. Social norms and politeness, for example, stand for the socio pragmatic knowledge; while considering various linguistic resources to produce a speech act are pragma-linguistic understandings.

Nevertheless, the point is that the issue of pragmatic assessment is relatively new. Pragmatic assessment mainly focuses on contextual relevance of L2 learner's language use (Oller, 1979). Several studies carried out in developing pragmatic tests. Brown (2001) used six types of tests. The instrument used was: Written Discourse Completion Tasks (WDCT), Multiple-choice Discourse Completion Tasks (MDCT), Oral Discourse Completion Tasks (ODCT), Discourse Role Play Tasks (DRPT), Discourse self-Assessment Tasks (DSAT), and Role-Play self-assessments (RPSA).

In the domain of socio pragmatics, Hudson, Detmer, and Brown (1992) used different methods in testing politeness and degree of directness of learners' apology, and request and Refusal competences. They developed six types of tests: oral DCTS, written DCTS, multiple choice DCT, role plays and self-assessment. In the domain of pragma linguistic issues of language, Roever (2005-2006) developed Web-based test of ESL pragmatics. Roever's instrument was more appropriate for both Asian and European test taker. Finally, Bachman (1990) claims pragmatic performances can be assessed through either rating on scales or counting the correct responses.

A new branch of pragmatic assessment is related to the issue of rating and rater criteria. Current research on pragmatic assessment focuses on the importance of rating criteria. Taguchi (2011) explored the rater variation in the assessment of speech act of request. He center on issues, such as "politeness markers", "amount of speech", "strategies", "directness" responses of EFL learners. Also, he found out that native raters had some inconsistencies in their rating. Lee (2012) studied rating behavior between

Korean and native English-speaking raters (NES) in which the study showed that Korean raters were more serious in scoring to linguistic component (grammar, sentence structure), while the NES raters emphasized the content and total scores.

Moreover, Walter (2007) investigated rater variation in pragmatic assessment. In his study, 42 learners of English attended a baseline activity with a native English-speaking tester for 10 to 15 minutes. Three oral pragmatic prompts were in the activity (an assessment, a compliment, and a pre-sequence) were embedded within three larger topic discussion. After each topic discussion, the prompts were delivered unconsciously. Two raters, a native and non-native speaker of English, assessed the baselines based on a four-point holistic rating scale. Dialogues between the raters were recorded when they determined differences in scoring and then analyzed. The results revealed that raters analyzed examinees' performance differently which caused different scoring. For instance, in the pragmatic target of compliment responses, the high proficient speaker emphasized his knowledge of normative patterns of compliment and compliment response in American English, while the low proficient speaker considered L1 transfer as possible source of non-normative compliment response. Examinee's fluency and clear pronunciation influenced low proficient speaker which led to a higher score.

Moreover, there are a number of studies in other area of assessment that investigated raters' variations of perspectives orientations (e.g., Brown 2000, 2003, 2005; Ducassee & Brown 2009; Johnson & Lim 2009; May 2006, 2009; McNamara & Lumely 1997; O'Loughlin 1996; Polit & Murray 1996). Using reflective verbal etiquette, these studies examined how raters' characteristics, gender, language background, experience, and competence affected their evaluation of L2 oral interviews, writing samples, and paired dyads. After rating learner's performance, raters were asked to summarize their reasons for awarding the

rating. Aspects of learners' performance that raters focused on a review of verbal reports revealed (e.g., linguistic features, discourse management, rhetorical organization, and listening behaviors). There is general agreement from the body of literature. This general consensus is that raters bring their own values and criteria in assessment, after training and they cohere to both criterion and non-criterion features.

In a more recent study, Tajeddin and Alemi (2013) focused on the assessment of EFL learners' pragmatic competence by native English raters' criteria. They investigated the criteria of native and non-native English teacher for the speech act of apology in L2. Their analysis of raters' remark manifested five criteria such as: apology, situation explanation, politeness, irrelevancy of Speech act, statements of alternatives.

For several years, great effort has been devoted to the study of request in speech acts. According to Trosborg (1995), request occurs when the speaker states his (her) wants to the hearer and want him (her) to do something for his benefits. In making a request, the speaker infringes on the recipient's freedom from imposition. The recipient may feel that the request is an intrusion on his/her freedom of action or even a power play (Blum-Kulka, House, & Kasper, 1989). In recent years, request is analyzed in forms of cross-cultural and interlanguage studies. Some researchers indicated the developmental pragmatics by comparing data from L2 learners at different levels of proficiency (Francis, 1997; Paren, 2002). Other researchers just compared nonnative and native speakers (Roose, 2000, p. 29). Several studies have revealed that differences in performing and realizing speech act demand teaching and testing it for EFL learners. These studies demonstrate the existence of differences in performing and realizing request speech act necessitate teaching and testing it for EFL learners. Thus, more studies are necessary in order to inform EFL teachers about different aspects of request speech act in teaching and

assessment processes. In view of this need in the literature, this study aimed at investigating what criteria inform non-native English speaking teachers' rating criteria with a focus on the speech act of request.

## METHOD

Forty non-native Iranian English teachers in the 25-35 age range were selected. The group of non-native English speaking raters (NNESRS) included English teachers from different languages institutes in Isfahan with various teaching experiences (classified into two levels of 1-5 and 6-11). These teachers were M.A holders or M.A students of TEFL (Teaching English as a Foreign Language). Therefore, they were familiar with the concept of L2 pragmatics and language testing.

In order to make sure in objective terms that the teachers were truly homogenous with regard to their English proficiency level, an Oxford Placement Test (OPT) was given to them. After obtaining the OPT results, it was decided to divide those teachers into high and low proficiency level. This being so, 40 teachers who met this homogeneity criterion were selected and assigned to two group (high and low proficient) involved in the study (20 teachers each). In addition, ten Iranian EFL learners from intermediate levels were selected for administrating the DCT.

The study was accomplished through 3 phases. The first phase was selecting 40 non-native English speaking teachers from language institutes in Iran with various teaching experiences and ten Iranian EFL learners from Intermediate groups. In the second phase, (OPT) test as a placement test was used to check the subjects' English proficiency. It was administrated to divide teachers into two groups of high and low level of proficiency. In the third phase, a DCT test was employed based on various degrees of formality, power, and distance to test the request speech act performance of EFL learners. Finally, both high and low proficient teachers were asked to assess the learner's performance with DCT test. For the

sake of convenience in rating, Likert scale is placed after every response for raters. Through the qualitative phase of data analysis, the Criteria noted by NNEERS were analyzed and categorized. Thus, the frequency of each criterion was calculated through quantitative procedures in order to find the dominant criteria.

Descriptive analyses including Mean and SD reveal the pragmatic assessment of teachers. NNEERS analyzed and categorized the criteria through the qualitative way of data analysis. Therefore, the frequency of each criterion is calculated through quantitative procedure in order to find the dominant criteria. Furthermore, t-test and chi-square were used to determine if there are any significant differences between NNEERS' rating scores and rating criteria.

## **RESULTS AND DISCUSSION**

NNEERS use various criteria during pragmatic assessment of EFL learners' requests. The following criteria were found in raters' comments:

- 1) *Directness*: This request criteria deals with the directness and indirectness of the EFL learners' productions. Example: "Leave me alone." An indirectness example is "It is too hot."
- 2) *Politeness*: This criterion refers to the degree of politeness of the EFL learners' request. Example: I think it's not very polite. The teachers should respect the students, especially in front of other students.
- 3) *Language usage accuracy*: This criterion is mainly about the accuracy of the structures, grammar, and lexical items of the produced sentences. Example: There are some grammar mistakes, such as "it" should be replaced by its reference "the book".
- 4) *Authenticity and cultural errors*: This criterion reflects the naturalness of the produced responses, as well as their cultural appropriateness regarding L2 society. Example: This sentence seems odd and unnatural. English people would never say that, especially the "go ahead" part.
- 5) *Style and register*: This criterion refers to the use of formal or informal style. Example: Asking your teacher informally might lead to misunderstanding.
- 6) *Explanation*: This criterion refers to brief explanation or introduction before making request. Example: I think it's better to add an introduction and clarify the request.
- 7) *Statement of optimal example*: This criterion refers to supplying various examples of the ideal request for the specified situations. Example: She/he could say: "I need that doll for my daughter. Please let me borrow it for a few days if you don't need it".
- 8) *Query preparatory and softeners*: This criterion refers to the importance of the use of preparatory expressions, such as could you, would you, etc., as well as words or phrases which can moderate the request (i.e. please, thank you, if it's OK with you). Example: "Pardon me" followed by the word "excuse me" is more favored.
- 9) *Conversers' relationship*: This criterion refers to social relationships, as realizing, establishing, sustaining, and changing social relations. Example: It depends on the closeness of the relationship. If it is an employee boss relationship, then the sentences are informal and not proper for this situation, whereas, it is considered proper between 2 friends.

Table 1. *Frequencies of responses to the DCT items by high proficiency teachers*

Item	Directness	Politeness	Accuracy	Authenticity	Style	Explanation	Statement of optimal	Query preparatory & softness	Conversers relationship	Expected Frequency	Chi-Square	df	Sig. (2-tailed)
1	4	6	6	9	2	8	8	8	9	6.70	6.90	8	.54
2	8	8	8	6	8	8	2	10	8	7.30	5.44	8	.70
3	4	5	2	8	6	6	10	7	6	6.00	7.00	8	.53
4	8	10	5	9	8	5	4	10	9	7.60	5.58	8	.69
5	5	10	7	9	6	7	8	4	10	7.30	4.90	8	.76
6	1	4	2	7	8	6	1	8	8	5.00	14.80	8	.06
7	5	8	9	9	8	8	8	10	8	8.10	1.83	8	.98
8	2	8	3	6	6	6	5	6	4	5.10	5.26	8	.72
<b>Total</b>	37	59	42	63	52	54	46	63	62	53.10	14.02	8	.08

Regarding the first DCT item and the frequencies it received from the high proficiency teachers, there were no statistically significant differences among the nine assessment criteria in as much as the  $p$  value under the Sig. (2-tailed) column for this item equaled .54. In fact, a  $p$  value less than the significance level (i.e.,  $p < .05$ ) signifies a significant difference, whereas a  $p$  value above .05 indicates that the difference failed to reach statistical significance.

The highest frequencies belonged to the criteria of authenticity ( $f = 63$ ), query of preparatory and softness ( $f = 63$ ), and conversers' relationship ( $f = 62$ ), whereas the lowest frequencies were those of directness ( $f = 37$ ), accuracy ( $f = 42$ ), and statement of optional example ( $f = 46$ ). However, as it was observed in Table 1, the differences among the nine assessment criteria by high-proficiency teachers did not reach statistical significance.

Table 2. *Frequencies of responses to the DCT items by low proficiency teachers*

Item	Directness	Politeness	Accuracy	Authenticity	Style	Explanation	Statement of optimal example	Query preparatory & softness	Conversers relationship	Expected Frequency	Chi-Square	df	Sig. (2-tailed)
1	6	0	10	0	6	2	8	6	8	6.60	5.73	6	.45
2	8	2	6	2	8	6	6	8	6	5.80	7.53	8	.48
3	2	6	2	2	8	4	6	4	2	4.00	10.00	8	.26
4	8	6	6	4	6	6	2	6	6	5.60	4.00	8	.85
5	6	6	8	6	8	4	4	3	6	5.70	4.23	8	.83
6	4	2	8	6	4	6	6	2	4	4.70	6.85	8	.55
7	4	6	8	6	8	6	8	4	4	6.00	4.00	8	.85
8	4	6	6	0	4	2	6	2	2	4.00	6.00	7	.54
<b>Total</b>	42	34	54	26	52	36	46	35	36	40.10	16.67	8	.03

For low-proficiency teachers, the Sig. values in the rightmost column of Table 2 show that for the eight individual items of the DCT, the differences among the nine assessment criteria were too small to reach statistical significance. However, adding up the frequencies of the criteria for all the DCT items yielded total frequencies for the assessment criteria, and the differences

among the total frequencies, as is shown in the lowest row of the table, reached statistical significance due to the fact that the  $p$  value in this row was found to be .03 ( $p < .05$ ). This means that, on the whole, low-proficiency teachers used significantly different criteria to assess the request speech act.

Table 3. Frequencies of responses to the DCT Items by high- and low- proficiency teachers

Item	Teachers	Directness	Politeness	Accuracy	Authenticity	Style	Explanation	Statement of optimal example	Query preparatory & softness	Conversers relationship	Chi-Square	Df	Sig. (2-tailed)
1	High	4	6	6	9	2	8	8	8	9	20.85	8	.008
	Low	6	0	10	0	6	2	8	6	8			
2	High	8	8	8	6	8	8	2	10	8	7.11	8	.52
	Low	8	2	6	2	8	6	6	8	6			
3	High	4	5	2	8	6	6	10	7	6	5.48	8	.70
	Low	2	6	2	2	8	4	6	4	2			
4	High	8	10	5	9	8	5	4	10	9	2.98	8	.93
	Low	8	6	6	4	6	6	2	6	6			
5	High	5	10	7	9	6	7	8	4	10	3.47	8	.90
	Low	6	6	8	6	8	4	4	3	6			
6	High	1	4	2	7	8	6	1	8	8	15.89	8	.04
	Low	4	2	8	6	4	6	6	2	4			
7	High	5	8	9	9	8	8	8	10	8	2.45	8	.96
	Low	4	6	8	6	8	6	8	4	4			
8	High	2	8	3	6	6	6	5	6	4	10.95	8	.20
	Low	4	6	6	0	4	2	6	2	2			
Total	High	37	59	42	63	52	54	46	63	62	26.61	8	.001
	Low	42	34	54	26	52	36	46	35	36			

The results presented in Table 3 revealed that high-and low-proficiency teachers differed significantly with respect to the criteria they used to assess the first item in the DCT ( $p=.008$ ). This might have occurred because of the fact that high-proficiency teachers and low-proficiency teachers differed with respect to the politeness (6 vs. 0), authenticity (9 vs. 0) and explanation (8 vs. 2) criteria.

In addition, high- and low-proficiency teachers were found to be significantly different in relation to their responses to the 6<sup>th</sup> DCT item ( $p = .04$ ). This difference might have been caused by the difference between the high- and low-proficiency teachers' responses to the accuracy (2 vs. 8), statement of optional example (1 vs. 6), and query preparatory and softness (8 vs. 2) criteria.

More importantly, the high- and low-proficiency teachers differed significantly in terms of the total frequencies of their responses to the DCT items (i.e., in terms of their overall assessment criteria) since the  $p$  value in the bottom row of the table was found to be .001. The high- and low-

proficiency teachers were probable significantly different in terms of their responses to the politeness (59 vs. 34), authenticity (63 vs. 26), query preparatory and softness (63 vs. 35), and conversers' relationship (62 vs. 36) criteria.

Although EFL teachers' rating criteria and patterns in ILP assessment has great impact on the process of teaching and testing of second language, this issue has remained understudied. The study examined whether higher and lower proficient teachers differed in assessing EFL learners' request, and what features of EFL learners' requests NNESRS used during pragmatic assessment.

The primary objective of this study was to explore the criteria employed by NNESRS in request rating process. The criteria were both Socio-pragmatic and Pragma-linguistic categories among both high and low proficient teachers. As an example, criteria like linguistic appropriacy or query preparatory and softeners belong to pragmatic aspect of language, while formality and social relationship or politeness fit into Socio-pragmatic category. According to Rasekh (2008) and Roever

(2007), the importance of both aspects in either teaching or testing pragmatic knowledge in previous researches.

Regarding the scoring of EFL learners' production, Iranian NNESRS acted differently, as their minimum and maximum scores in most situations. The deficiencies in high and low proficient raters' evaluation and the significant differences between them emphasize the inadequacy of some of the low proficient teachers' pragmatic knowledge and the necessity of developing teacher training courses, especially pragmatic training for NNES teachers (Alemi, 2012; Harlig & Hartford, 1997; Rasekh, 2005; Rose, 2005).

For the third questions, quasi-experimental design producer was carried out. Based on the achieved results, there was significant difference between high and low proficiency teachers regarding pragmatic assessment of speech act request. Teachers in the high proficiency group could more successfully enhance pragmatic competence compared to the low proficiency group in pragmatic assessment.

## CONCLUSION

The present research was an attempt in which pragmatic assessment considered to be the ideal method for pragmatic assessment of speech act request. So, the homogenized participants were divided into two groups of high and low proficiency. In this study, nine different criteria are employed by NNESRs in rating request productions. High proficient teachers apply certain criteria in evaluating the request speech act of EFL learners more frequently. These criteria include Authenticity, Politeness, Query preparatory and softness, Conversers relationship, while low proficient teachers use some criteria such as Directness, Accuracy, Style, and Statement of optimal example, more frequently. These differences could be due to lack of pragmatic knowledge on the part of NNESRs in which there is cultural difference between L1 and L2 causing pragmatic misunderstanding.

The study also indicates the important of L2 pragmatic and the need for pragmatically appropriate learning materials. In countries like Iran, teachers and learners do not have any easy access to native speakers or authentic learning materials. In fact, learners need pragmatic instruction as a part of their language education while most of the textbooks for language learning lack sufficient L2 pragmatic exercises and do not consider cross-cultural differences between L1 and L2 societies (Alemi & Irandoost, 2012; Alemi, Roodi, & Bemani, 2013; Safa, Moradi, & Hamzavia, 2015).

Finally, it is hoped that research in L2 pragmatics will not only improve our understanding of pragmatic development in speech act realization and of the nature of strategies, but further studies will be done to find EFL teachers' rating criteria in assessing understudied speech acts such as criticism, congratulation, etc.

## REFERENCES

- Alemi, M., Eslami-Rasekh, Z., & Rezanejad, A. (2015). Iranian non-native English speaking teachers' rating criteria regarding the speech act of compliment: An investigation of teachers' variables. *Journal of Teaching Language Skills*, 6(3), 21-49.
- Byon, A. S. (2004). Socio-pragmatic analysis of Korean requests: Pedagogical setting. *Journal of Pragmatics*, 36, 1973-1704.
- Chen, S. (2007). Interlanguage requests: A cross-cultural study of English and Chinese. *The Linguistics Journal* 2(2), 32-52.
- Cohen, A. D. (2014). Towards increased classroom assessment of pragmatic ability. *Iranian Journal of Language Testing*, 4(1), 4-25. Retrieved from <http://ijlt.ir/journal/images/PDF/420-2014-4-1.pdf>.
- Jalilifar, A., Hashemian, M., & Tabatabaee, M. (2011). A cross-sectional study of Iranian EFL learners' request strategies. *Journal of Language Teaching and Research*, 2(4), 790-803. doi:10.4304/jltr.2.4.790-803.
- Koosha, B., & Dastjerdi, H. V. (2012). Investigating pragmatic competence: The case of requests in interchange series. *Asian Social Science*, 8(1), 45-61.
- Plough, I. C., Briggs, S. L., & Van Bonn, S. (2010). A multi-method analysis of evaluation criteria used to assess the speaking proficiency of graduate student instructors. *Language Testing*, 27(2), 235-260.

- Rezanejad, A. (2013). *Rating EFL learner's interlanguage pragmatic competence by native and non-native English speaking teachers* (Unpublished master's thesis). Sharif University of Technology, Tehran, Iran.
- Roever, C. (2010). Effects of native language in a test of ESL pragmatics: A DIF approach. In G. Kasper, H. Thi Nguyen, D. R. Yoshimi, & J. Yoshioka (Eds.), *Pragmatics and language learning* (Vol. 12, pp. 187–212). Honolulu, HI: National Foreign Language Resource Center.
- Safa, M. A., Moradi, M., & Hamzavia, R. (2015). Iranian EFL teachers and learners perspective on potentiality of Top Notch series for intercultural competence development. *Iranian Journal of Language Teaching Research*, 3(2), 47-66. Retrieved from [http://www.urmia.ac.ir/sites/www.urmia.ac.ir/files/\(4\).pdf](http://www.urmia.ac.ir/sites/www.urmia.ac.ir/files/(4).pdf).
- Tajeddin, Z. (2014). *Interlanguage pragmatics and language teaching*. Tehran, Iran: Jungle Publications.
- Trosborg, A. (1995). *Interlanguage pragmatics: Requests, complaints, and apologies*. Berlin: Mouton de Gruyter. doi:10.1515/9783110885286.
- Walters, F. S. (2004). *An application of conversation analysis to the development of a test of second language pragmatic competence*. Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign.
- Walters, F. S. (2007). A conversation-analytic hermeneutic rating protocol to assess L2 oral pragmatic competence. *Language Testing*, 24(2), 155–183.
- Walters, F. S. (2009). A conversation analysis-informed test of L2 aural pragmatic comprehensions. *TESOL Quarterly*, 43(1), 29–54.
- Woodfield, H. (2008). Interlanguage requests in English: A contrastive study. In M. Pütz & J. N. V. Aertselaer (Eds.), *Contrastive pragmatics: Interlanguage and Cross-cultural Perspectives* (pp. 231-264). Berlin/New York: Mouton De Gruyter.
- Youssef, A. M. (2012). Study of request strategies employed by Libyan and Malay postgraduate students at USM. *International Journal of Learning & Development*, 2(2), 144-151.